

Statistical Multiplexed Optical Burst-Train Switching Network

Oliver Yu, Ming Liao

Department of Electrical and Computer Engineering,
University of Illinois at Chicago, Chicago, IL 60607

Abstract—The Optical Burst-Train Switching (OBTS) network efficiently transports multi-granular sub-wavelength constant-bit-rate (CBR) and variable-bit-rate (VBR) traffic streams over an all-optical network with sparse wavelength convertibility while minimizing access blocking of traffic streams. Sub-wavelength traffic streams are shaped into periodic burst-trains at network edge with variable burst sizes. Proactive periodic reservation of burst-trains minimizes reservation signaling and scheduling overheads. OBTS combines variable-sized-slot time division multiplexing over network core and stream-based statistical multiplexing at network edge to provide guaranteed bandwidth for CBR traffic, while minimizing data loss of average-provisioned VBR traffic. Performance of OBTS is analyzed in terms of stream or burst-train blocking and data loss of average-rate provisioned VBR traffic. The tuning effects on OBTS performance are investigated with regards to wavelength convertibility factor, switching latency, periodic scheduling frame size, burst size and traffic burstiness.

I. INTRODUCTION

Dynamic optical flow and burst switched networks based on wavelength division multiplexing (WDM) and all-optical switches have been increasingly deployed to bypass routers and electronic switches; so as to avoid disparity between optical transmission rate and electronic device processing rate, and to enable quality-of-service (QoS) performance in terms of throughput and delay directly in the optical layer. Dynamic wavelength-routed optical networks (WRON) employ wavelength or optical circuit switching (OCS), and two-way per-flow reservation to support broadband inelastic streaming traffic with multi-gigabit rate through a lightpath of wavelength-switched channels. However, OCS-based WRON would be inefficient to support burst traffic or narrowband inelastic streaming traffic with sub-wavelength bandwidth granularities.

Optical burst switching (OBS) networks ([1]-[5]) are designed to provide efficient transport of sub-wavelength burst traffic. These schemes effectively realize statistical multiplexing (SM) of optical wavelength resources through burst aggregation at network edge, and burst level transmission granularity with delayed reservation in network core. OBS schemes can employ one-way reservation per-

burst (JET [1] and JIT[2]) to reduce end-to-end burst transport delay, but this may cause potential data loss due to burst contention. Two-way reservation per burst [3], or per sequence of periodic bursts [4] can be employed to avoid burst contention, but data loss still occurs due to burst reservation blocking [6]. Prioritized QoS schemes [5] can be realized by assigning differential extra offset time to bursts.

Time division multiplexing (TDM) over WDM networks [7][8] enables constant-bit-rate (CBR) inelastic sub-wavelength streaming traffic flows to time-share a lightpath efficiently, through coordinated traffic controls at network edge and core. At the network edge, each sub-wavelength CBR traffic flow is shaped into a sequence of periodic bursts with wavelength granularity. In the network core, each wavelength channel between all-optical switches is time-shared by multiple burst-sequenced traffic flows. Within each all-optical switch, the cross-connection of each incoming-outgoing port pair is dynamically configured with periodic fixed-sized active time slots, during which bursts of shaped traffic flows are allowed to hold the corresponding optical wavelength channel. The active time slots are separated by idle intervals, during which transmission over the corresponding optical wavelength channel is stopped due to switch latency or cross-connection setup delay.

The control plane of TDM/WDM switched network maintains a periodic burst scheduling frame, which consists of a number of periodic active time slots. For each burst-sequenced traffic flow, the traffic controller would schedule each burst to transmit in one or more active time slots (contiguous or non-contiguous) with each successive scheduling frame. This allows periodic scheduling and reduces complexity of scheduling control database.

Existing TDM/WDM schemes employ time-wavelength-space-router (TWSR) [7] in the core, which require buffering or delaying of burst when there is a synchronization mismatch between burst arrival/departure times and switch ON/OFF times. Due to immature optical buffer technology, programmable fiber delay lines (FDLs) are employed instead for time re-synchronization. Consisting of a series of all-optical switching fabrics and optical fibers, FDLs add significant implementation costs to these TDM/WDM schemes.

Multi-granular stream optical burst switching (MGS-OBS) [9] extended the TDM/WDM scheme with variable-

sized control time slots to support sub-wavelength streaming traffic of diverse bandwidth granularities. This also allows flexible burst scheduling and optimizes wavelength utilization by minimizing the idle intervals of optical wavelength channels.

This paper presents the Optical Burst-train Switching (OBTS) scheme, which further extended the TDM/WDM scheme to efficiently support both CBR and variable-bit-rate (VBR) traffic by combining stream-based statistical multiplexing at network edge and periodically scheduled variable-sized-slot TDM over network core. This enables the provision of guaranteed bandwidth for CBR traffic, while minimizing data loss of average-provisioned VBR traffic. Furthermore, OBTS does not require the support of FDLs as it employs guard-time between bursts to handle the mismatch between burst arrival/departure times and switch ON/OFF times.

The rest of this paper is organized as follows. Section II presents the OBTS architecture. Section III shows the simulation results, and section IV concludes the paper.

II. OPTICAL BURST-TRAIN SWITCHING

At the network edge, OBTS shapes sub-wavelength traffic streams into periodic light-rate burst-trains. In the network core, these burst-trains are transported via variable-sized slotted-time lightpaths established by proactive periodic reservation to minimize signaling overhead and scheduling complexity.

The slotted-time lightpaths or sub-wavelength optical circuits are established dynamically with variable bandwidth capacities to support on-demand multi-granular streaming

traffic requests with guaranteed bandwidth and data delivery requirements. As illustrated in Fig. 1, the OBTS architecture specifies the edge and core controls to enable scheduling and switching of the burst-trains. The edge node controller (ENC) reserves the burst-train for an admitted stream and shapes the buffered data into bursts. It consists of three major components: (1) the traffic conditioner that shapes the buffered data into bursts and performs edge data statistical multiplexing; (2) the burst-train scheduler that allocate bursts to corresponding time slots; and (3) the resource reservation signaling controller that interacts with the core network controller to provide the signaling functionality.

The core node controller (CNC) consists of two components: (1) the resource reservation signaling controller that reserves slotted-time lightpaths in the network core in response to requests from the corresponding peer in the edge node controller; and (2) the periodic burst scheduler to dynamically configure the optical switches for each burst-train.

The stream access request specifies the destination and the bandwidth requirement of the streaming traffic. The ENC burst-train scheduler executes traffic admission control, and attempts to setup a sub-wavelength connection by reserving wavelength and time slot resources in the core. Connection setup includes route selection, resource discovery and reservation, and switch configuration.

The ENC burst-train scheduler invokes the resource reservation signaling controller to discover resource availability of wavelengths and time slots. Based on the employed routing-wavelength-time assignment (RWTA) algorithm, the scheduler selects the route, wavelength channels and time slots of the slotted-time lightpath to transport the burst-train. Then the scheduler invokes the signaling controller to reserve the wavelength and time slot resources along the selected route.

The ENC and DNC resource reservation signaling controllers interact via the robust and fast optical reservation protocol (RFORP) [10], which extends the RSVP protocol. The CNC periodic burst scheduler executes switch configuration along the selected route.

After successful establishment of the slotted-time lightpath, the ENC burst-train scheduler invokes the traffic conditioner that triggers burst-train transmission. The streaming data is allowed to enter the edge aggregation buffers. The traffic conditioner shapes the buffered data into burst-trains, and executes stream-based statistical multiplexing for VBR traffic to optimize wavelength utilization and minimize data loss due to potential mismatch between average rate provisioning and peak rate arrival. The core all-optical switches cross-connect and transport the bursts based on the configured reservation database.

Sub-sections 2.A - 2.C presents detailed designs of the periodic burst scheduler, the burst-train scheduler, the traffic controller, and the resource reservation signaling controller.

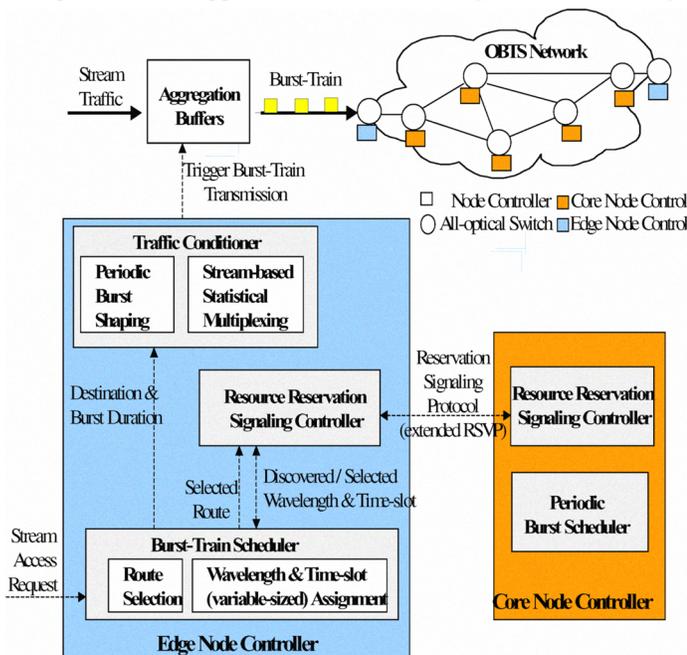


Fig. 1 OBTS System Architecture

A. Periodic Burst Scheduler

To reduce the signaling overhead and RWTA computational complexity, the periodic bursts of the burst-train are reserved once during the connection setup phase. The time scale is divided into fixed-sized time frames (scheduling frame). In each time frame, each admitted request will be supported by variable-sized time-slot(s) to accommodate multi-granular bandwidth requirement while minimize the incurred switching latency overhead. The periodic burst scheduler in core node records the switch ON/OFF time for each burst-train in its database. These switching ON/OFF operations are periodically assigned so that the number of time-slots for each burst-train in each frame is the same; and the relative position of the time-slots reserved for each burst-train are the same in each frame. The aggregated data in each buffer will be shaped into a periodic burst-train accordingly. The bursts are then periodically dumped to the core network based on the assigned wavelength and time-slots. Since the ending time of a traffic stream may not be available at the time of reservation, the source needs to signal the core node controllers to explicitly release periodic burst-train reservation for the traffic stream.

The advantages of introducing periodic scheduling frame are as follows: (i) switch controllers only need to maintain the relative switch ON/OFF time with respect to frame beginning time; (ii) only the switch ON/OFF time within one time frame needs to be maintained, rather than maintaining it for the entire time duration that the stream is active. Each burst is associated with a fixed switching latency overhead. To reduce the total switching latency overhead, the number of time-slot reserved for each stream in each time frame should be minimized.

As illustrated in Fig. 2 (where T refers to the length of time frame), OBTS may reserve more than one burst per frame. For example, traffic A reserves two bursts per frame, while traffic B reserves one burst per frame. Each burst is associated with a switching latency overhead t_{sw} , which represents the time all-optical switch needed to do the switching. The actual propagation time may vary from its given value due to the temperature change, a guard time t_G is set for each burst to handle the mismatch between burst arrival/departure times and switch ON/OFF times.

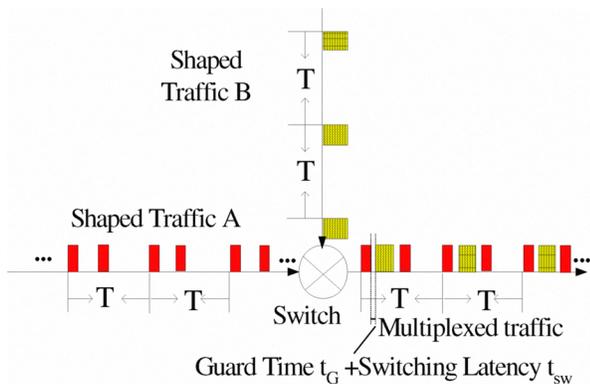


Fig. 2 Periodical Scheduling and Time Division Multiplexing

B. Burst-train Scheduler and RWTA

To support sub-wavelength traffic streams, the burst-train scheduler needs to determine the end-to-end routing, as well as the wavelength and time-slot assignment on each link for a connection request. Such routing, wavelength and time-slot assignment problem is called RWTA problem. The objective of a generic RWTA problem is to maximize the throughput (for static traffic pattern) or minimize the burst-train blocking (for dynamic traffic pattern). To reduce the computational complexity, the RWTA for OBTS is decoupled into two parts: route selection with shortest-hop routing; and combined wavelength and time assignment that is described as follows.

Notations:

- $L_{i,j}$ the j th link of the route for the i th traffic request
- r_i the required bandwidth (bits/sec) for the i th request.
- C the bandwidth (bits/sec) of each wavelength.
- T the length (sec) of a scheduling frame.
- $S_{i,j}$ the starting time of the variable time-slot reserved for the i th request on $L_{i,j}$.
- $E_{i,j}$ the ending time of the variable time-slot reserved for the i th request on $L_{i,j}$.
- $\lambda_{i,j}$ the selected wavelength on $L_{i,j}$.

The optimization objective is to find a set $(S_{i,j}, E_{i,j}, \lambda_{i,j})$ for each $L_{i,j}$, to maximize total throughput:

$$\text{Maximize: } \sum_i E_{i,1} - S_{i,1}$$

Constraints:

- Wavelength continuity constraint:
 $\lambda_{i,j} = \lambda_{i,j+1}$ unless the node between $L_{i,j}$ and $L_{i,j+1}$ is capable of wavelength conversion.
- Propagation delay constraint:
 $S_{i,j+1} - S_{i,j} = E_{i,j+1} - E_{i,j} = \text{propagation delay on } L_{i,j}$; unless the node between $L_{i,j}$ and $L_{i,j+1}$ is equipped with programmable FDL. (Note in this paper, we assume that there is no programmable FDL.)
- Variable time-slot size constraint:
 $E_{i,j} - S_{i,j} = r_i * T / C$ for all j , if the stream request i can be supported; otherwise $E_{i,j} - S_{i,j} = 0$, for all j . This means that the size of reserved variable time-slot should be large enough to satisfy the request bandwidth.
- Time-wavelength contention constraint:
 $(S_{i,j}, E_{i,j}) \cap (S_{i',j'}, E_{i',j'}) = \emptyset$ or $\lambda_{i,j} \neq \lambda_{i',j'}$, where active duration of stream i and i' overlaps, and $L_{i,j}$ and $L_{i',j'}$ are referred to the same link.

To theoretically solve this optimization problem, the entire traffic matrix including the information of future arrival must be given. In realistic systems, such information may not be pre-known since most of the connection requests are dynamic on-demand traffic. A heuristic is proposed to

provide a sub-optimal solution for the WTA problem. The basic idea is to minimize the blocking probability for each individual incoming connection request.

The WTA problem is solved by the earliest time-slot search, which executes exhaustive searches for all the common available wavelengths and time assignments along the selected route, subject to the constraints described previously. The basic procedures are listed as follows:

- It will first search for all wavelengths, and find all available time-slots that are large enough to support the bandwidth requirement of the request.
- If multiple slots are found, it will compare the different available time-slots among all the wavelengths, and use the earliest available time-slot relative to the beginning of time frame.
- The effect of this algorithm is to shift the time assignment on first link to the beginning of time frame. Similar to the void filling algorithm in OBS scheduling, this will make room to support future stream request.

C. Traffic Conditioner

The traffic conditioner shapes the edge buffered data into periodic burst train according to the RWTA decision made in the burst scheduler. Let the bandwidth requirement be R , the capacity per wavelength be C , the length of periodic time frame be T . If only one burst is reserved in each frame, the burst size is determined by:

$$t_b = T \cdot R / C \quad (1)$$

If the streaming traffic generates packets in CBR pattern, the traffic conditioner will simply shape the aggregated data into a periodic burst-train according to the constant time slot size allocation. There will be no dropping of packets at the edge. When VBR streaming traffic is shaped and allocated with time slot size according to the average bit rate requirement, the actual data aggregation size in a particular scheduling frame may differ from the reserved bandwidth. To reduce the data loss due to the aggregated data size exceeding the reserved burst size, multiple VBR traffic streams with the same source and destination can share their reserved bandwidth through edge statistical multiplexing. This can be easily achieved by putting incoming data into different edge buffers according to the destination nodes. The packets sharing same end-to-end path in core network will be stored in the same edge buffer.

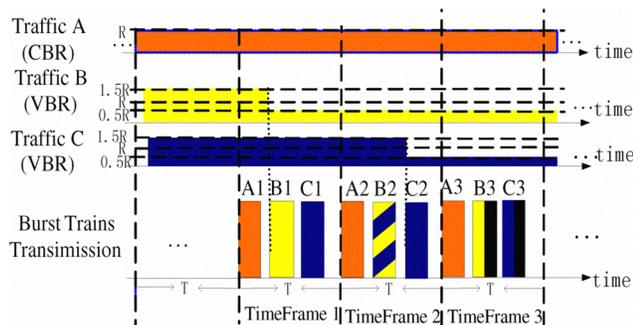


Fig. 3 TDM and Stream-based Statistical multiplexing of OBTS

Fig. 3 gives a simple example of how OBTS schedules and allocates bandwidth resource for burst-trains of sub-wavelength CBR and VBR streaming traffic. As illustrated in the figure, three streaming traffic are sharing one wavelength along the same route. Traffic A is a CBR stream with constant data rate of R ; traffic B and C are VBR streams with a maximum data rate of $1.5R$.

Firstly, the burst-trains for traffic A , B and C are reserved with the same capacity $R \cdot T$, which is equal to the amount of data aggregated in one time frame with the average data incoming rate of R . Traffic A has a constant data rate of R , so its reserved capacity perfectly matches the aggregated data size. Traffic B and C have different data rate in these three time frames, so there is mismatch between average provisioning rate and actual data arrival rate.

Secondly, stream-based statistical multiplexing of VBR traffic B and C is applied to further improve wavelength utilization. In time frame 1, the aggregated data of both traffic B and traffic C exceeds its reserved burst capacity and results in edge data dropping. In time frame 2, the aggregated data of traffic B does not use up its reserved burst capacity since its data rate is less than R . The unused capacity for traffic B will be allocated to the aggregated data from traffic C . In time frame 3, neither traffic B nor traffic C use up the corresponding reserved capacity.

III. RESULTS

To analyze exact effects of different system parameters on the performance of the proposed OBTS, simulations are conducted based on the typical 14-node NSF network topology (Fig. 4). The detailed effects of different system parameters (switching latency, wavelength convertibility, frame size) on burst-train blocking will be presented in Fig. 5-7. Fig. 8 will compare the performance of fixed-sized-slot TDM with variable-sized-slot TDM. Fig. 9 will explain the effect of stream-based edge statistical multiplexing.

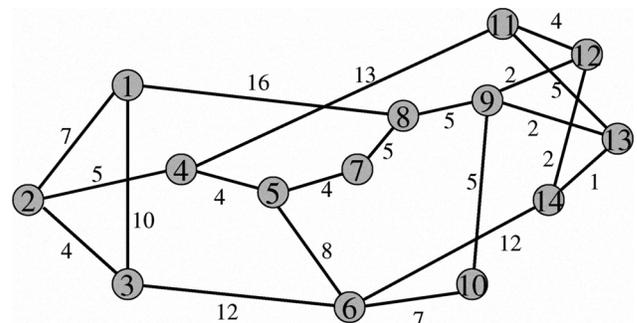


Fig. 4 14-node NSF network topology

The parameters of the topology model are as follows: number of nodes $N = 14$; number of links $|E| = 42$; number of wavelength per link $W = 4$. A generic time unit (TU) is used for flexibility. The propagation delay for each link is assumed to be 1 TU. The bandwidth capacity C of each wavelength = 10Mb/TU . The edge-to-edge connection path is pre-computed via the shortest hop algorithm. Each connection has the same bandwidth request $R = 125\text{Kb/TU}$.

It is assumed that fiber delay line is not available in the core network.

Simulation experiments are executed over a C++ based simulator, and 5000 traffic requests are simulated for each sample point in all the figures. Unless stated otherwise, the wavelength convertibility factor (ratio of the number of wavelength convertible nodes over total number of nodes) is set to be 3 out of 14; the length of time frame T is set to be 300 TU; switching delay is set to be 0.3 TU.

The normalized per-wavelength traffic loading is given by:

$$PerWavelengthLoading = \sum_i \frac{\alpha_i \cdot H_i}{\beta_i \cdot |E| \cdot W} \cdot \frac{R_i}{C} \quad (2)$$

where i is the index of end-to-end path, α_i is the average connection arrival rate on the i_{th} path, β_i is the average connection service rate on the i_{th} path, H_i is the number of hops on the i_{th} path, R_i is the average bandwidth requirement of the streams along this path. The loading parameter shown in the following figures refers to this normalized per-wavelength loading.

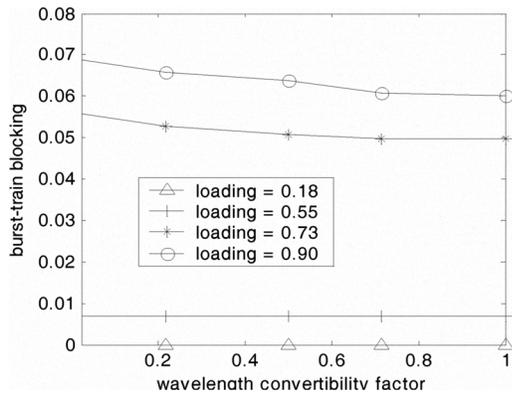


Fig. 5 Effect of Wavelength Convertibility Factor on Burst-Train Blocking under Different Loading

The burst-train blocking is composed of two kinds of components: dynamic contention blocking and congestion blocking. Dynamic contention blocking happens when two or more on-demand requests try to reserve dynamically the same wavelength and time duration on the same link. OBTS minimizes the contention probability by employing one-time proactive periodic reservation for the entire burst-train. On the other hand, congestion blocking happens when a request cannot find a time-wavelength assignment to satisfy the constraints given in section 2.B. Wavelength conversion relaxes the wavelength continuity constraint, which would reduce congestion blocking.

Fig. 5 presents the effect of wavelength convertibility factor on burst-train blocking under different loading conditions. When the loadings are low (0.18, 0.55), the congestion blocking component would be negligible and will not contribute to the overall burst-train blocking. Consequently, wavelength convertibility factor has no effect on burst-train blocking. When the loading are higher (0.73,

0.90) but still below saturation, adding more wavelength convertible nodes can only reduce the burst-train blocking very slightly. From network provider perspective, this result is desirable since OBTS can work effectively without costly wavelength converter when loading is below saturation.

Fig. 6 shows the effect of switching latency on burst-train blocking probability under different load settings. Switching latency is the time that the optical switch needs to connect the input port to an output port before any data can be transmitted between these two ports. Increasing switching latency will increase the data transmission interrupt time associated with each burst. This would increase the resource needed to support given traffic. Effectively this would result in higher blocking when total capacity is fixed. As it is shown in Fig. 6, burst-train blocking increases linearly as the switching latency increases, and this effect is more significant when loading is higher. For example, when load is higher than 0.36, increasing the switching latency from 0 to 3 TU can increase the blocking by 15~20%. Based on this result, the network designers can then find the best cost effective switching latency settings for their own topology.

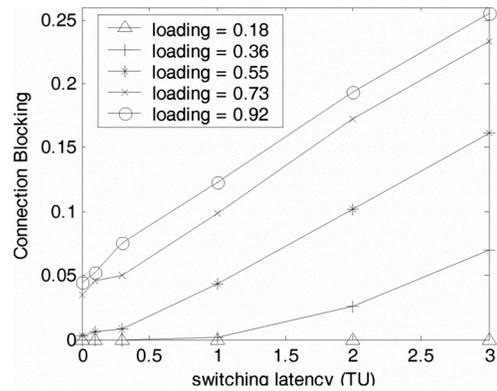


Fig. 6 Effect of Switching Latency on Burst-train blocking under Different Loading

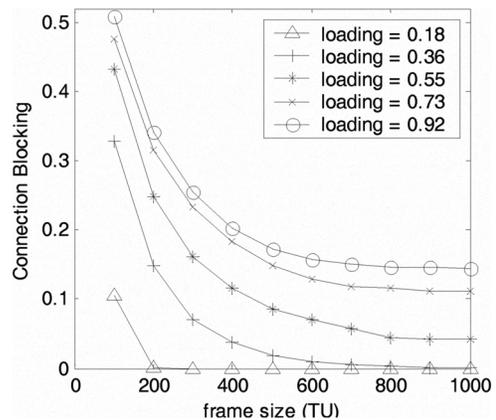


Fig. 7 Effect of Periodic Scheduling Frame Size on Burst-train Blocking

One way to reduce the blocking that is increased by switching latency is to enlarge the periodic scheduling frame size. According to (1), the ratio of reserved burst size to the

frame size equals to the ratio of requested bandwidth requirement to the total capacity per wavelength. When frame size increases, the total number of burst components per frame stay as a constant number. Therefore, increasing the frame size will increase the size of each burst and reduce the total number of bursts in a given time period. Since the overhead incurred by switching latency is a fixed value for each burst, this will effectively reduce the total switching latency overhead. Fig. 7 shows the effect of frame size on burst-train blocking. Switching latency for this figure is set at 3 TU to show that increasing the frame size can counteract the switching latency overhead.

Fig. 7 shows that increasing frame size can significantly reduce the burst-train blocking. However, this effect will become less effective if the frame size is already large. This is expected because the number of bursts in a give time period is inverse proportional to the frame size. In realistic network, increasing the frame size will require corresponding increase in edge data buffer size, which may have an upper limit. The network designer can find the best settings of the frame size for their own topology based on such graph.

The comparison between variable-sized-slot TDM and fixed-sized-slot TDM is presented in Fig. 8. As it is shown, variable-sized-slot TDM (OBTS) can achieve 3% lower blocking than fixed-sized-slot TDM when loading is higher than 0.50. The reason is that a burst is supported by one time-slot in variable-sized-slot TDM, but it may need to reserve several time-slots in fixed-sized-slot TDM. Since there is a fixed switching latency overhead associated with each time-slot, variable-sized-slot TDM can significantly reduce the overall switching latency overhead compared to that of fixed-sized-slot TDM.

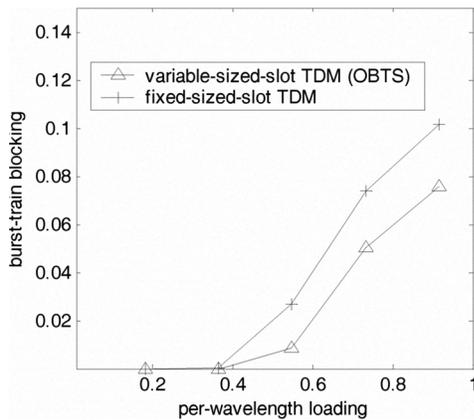


Fig. 8 Variable-sized-slot TDM versus Fixed-sized-slot TDM

Fig. 9 presents the performance of data loss rate versus burstiness of admitted traffic. Burstiness is defined as the ratio between the standard deviation and the mean value of the aggregated data size per burst. This result is obtained when the loading is around 0.35. The data loss rate is almost tripled when burstiness increases from 0.1 to 0.5. However, the data loss rate remains stable with statistical multiplexing.

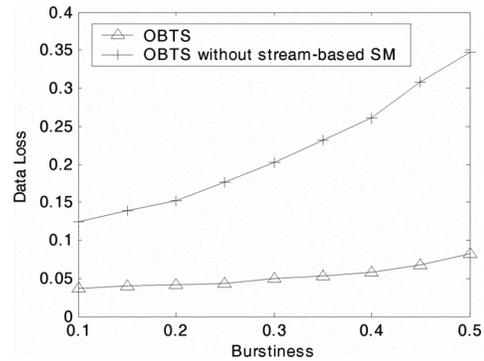


Fig. 9 Effect of Stream-based Statistical Multiplexing

IV. CONCLUSIONS

OBTS provides efficient transport of sub-wavelength CBR and VBR stream traffic over an all-optical wavelength switched network. It is achieved by shaping sub-wavelength traffic streams into periodic light-rate burst-trains, which are transported through variable-sized slotted-time lightpaths and edge stream-based statistical multiplexing to optimize wavelength utilization. Furthermore, OBTS employs proactive periodic burst reservation for each burst-train to minimize signaling overhead and scheduling complexity. The performance of OBTS is evaluated through simulation. It shows that the OBTS achieves lower burst-train reservation blocking and data loss rate by employing the variable-sized-slot and edge statistical multiplexing. The performance tuning results show that OBTS can provide satisfactory performance for the streaming type of traffic, regardless of the wavelength convertibility ratio in core network. It is also shown that the burst-train blocking can be greatly reduced by decreasing switching latency or increasing frame size.

REFERENCES

- [1] C. Qiao, and M. Yoo, "Optical burst switching (OBS)-A new paradigm for an optical internet," *J. High Speed Networks*, vol. 8, no. 1, pp. 69-84, 1999.
- [2] J. Wei, and R. McFarland, "Just-in-time signaling for WDM optical burst switching networks," *IEEE J. Lightwave Technology*, vol.18, no. 12, pp. 2019-2037, Dec. 2000
- [3] M. Duser, and P. Bayvel, "Analysis of a dynamically wavelength-routed optical burst switched network architecture," *IEEE J. Lightwave Technology*, vol. 20, no. 4, pp.574-585, 2002.
- [4] S. Sheeshia, and C. Qiao, "Synchronous optical burst switching," *Proc. BROADNETS 2004*, pp. 4-13, Oct. 2004.
- [5] M. Yoo, C. Qiao, S. Dixit, "QoS Performance of Optical Burst Switching in IP-over-WDM Networks," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 10, pp. 2062-2071, Oct. 2000.
- [6] A. Sridharan, and K. Sivarajan, "Blocking in all-optical networks," *IEEE/ACM Trans. On Net.*, vol 12, no. 2, pp.384-397, April 2004.
- [7] N. Huang, G. Liaw, and C. Wang, "A Novel All-Optical Transport Network with Time-Shared Wavelength Channels," *IEEE J. Select. Areas Commun.*, vol.18 no. 10, pp. 1863-1875, Oct. 2000.
- [8] B. Wen, and K. M. Sivalingam, "Routing, wavelength and time-slot-assignment algorithms for wavelength-routed optical WDM/TDM networks," *IEEE J. Lightwave Technology*, vol.23, no. 9, pp. 2598-2609, Sept. 2005.
- [9] O. Yu, M. Liao, Y. Cao, "Multi-Granular Stream Optical Burst Switching," in *Proc. BROADNETS 2006*, pp.1-10, Oct. 2006.
- [10] O. Yu, A. Li, Y. Cao, H. Xu, M. Liao, L. Yin, "Multi-domain Lambda Grid Data Portal for Collaborative Grid Applications," *Elsevier Journal of Future Generation Computer Systems*, vol. 22, issue 8, pp. 993-1003, Oct. 2006.